

A Comparative Analysis of Occupancy Data Collectors for the User Profile of the I-85 HOV-to-HOT Conversion

A Thesis
Presented to
The Academic Faculty

by

D. Kemp Edwards

In Partial Fulfillment
of the Requirements for the Research Option in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
August 2013

A Comparative Analysis of Occupancy Data Collectors for the User Profile of the I-85 HOV-to-HOT Conversion

Approved by:

Dr. Randall L. Guensler, Advisor
School of Civil & Environmental Engineering
Georgia Institute of Technology

Dr. Michael P. Hunter
School of Civil & Environmental Engineering
Georgia Institute of Technology

Dr. Donald R. Webster, Associate Chair
School of Civil & Environmental Engineering
Georgia Institute of Technology

Date Approved: August 2, 2013

Table of Contents

Chapter 1: Introduction	1
Chapter 2: Literature Review	2
2.1 History of HOV Lane Systems	2
2.2 Purpose of HOV/HOT Systems	3
2.3 Performance Measures of HOV/HOT Systems	5
2.3.1 Capacity and Flow	5
2.3.2 Occupancy.....	5
Chapter 3: Methodology	6
3.1 Data Collection Methods	6
3.1.1 License Plate Data Collection - Apparatus Configuration.....	7
3.1.2 Occupancy Data Collection	9
3.1.3 Double-count Occupancy Data Collection	12
3.1.4 License Plate Data Processing	12
3.2 Matching Occupancy Data with License Plate Data.....	14
Chapter 4: Results.....	18
Chapter 5: Discussion and Conclusion	23
5.1 Discussion of Results.....	23
5.2 Concluding Remarks.....	24

List of Tables

Table 1.0: Vehicle Body Type Re-Classification	11
Table 2.0: Matched License Plate Videos and Occupancy Data	20
Table 2.1: Analysis Data – Entry Totals by Type	20

Table 2.2: Analysis Data – Entry Percentages by Type	21
Table 2.3: Analysis Data – Effective Entries by Type	21

List of Figures

Figure 1: Example of configurations for the bridge cameras.....	7
Figure 2: Example of camera setup and configuration with highway view.....	8
Figure 3: Example of gore area camera and stake setup.	10
Figure 4: Keypad for Occupancy Data Collection Process.....	11
Figure 5: Sample screenshot of video processing program with a view of the HOT lane and adjacent lane.	13
Figure 6: Flow Chart of Matching Occupancy and License Plate Data.....	16
Figure 7: Sample screenshot of all three data streams matched and corrected	18
Figure 8: Total Effective Entries by Type	22
Figure 9: Total Error Rates by Type	22

Chapter 1: Introduction

Highway lane management has developed into a priority for transportation planning due to rising congestion levels, the environmental impact of vehicle emissions, and dwindling funding for transportation projects. High-occupancy vehicle (HOV) lanes were first introduced in 1969 to improve transit ridership in New Jersey [1]. High-occupancy toll (HOT) lanes were initially introduced in the 1990s to allow vehicles formerly prohibited from using HOV lanes to use the excess capacity still available in underutilized HOV lanes [2]. Typically, HOV and HOT lane systems allocate specific lanes for vehicles with two or more occupants or with certain emissions characteristics. HOV/HOT systems can be assessed by their effects on lane capacity, vehicle occupancy, travel time, reliability, safety, or other factors [3].

The metropolitan Atlanta interstate system is currently undergoing a conversion process from HOV to HOT lanes. To assess the effectiveness of the HOT lanes, the Georgia Department of Transportation (GDOT) requires accurate data about the interstate use. A Georgia Institute of Technology (Georgia Tech) research team is responsible for generating and analyzing a profile of I-85 users using manually collected vehicle occupancy and license plate data. The objective of this study is to compare the data collection performance of new and experienced research team data collectors by evaluating the occupancy data of the I-85 user profile.

The research team collects data quarterly for I-85 users using simultaneous manual and video collection techniques during two-hour peak morning and afternoon traffic periods. The data sets are matched and corrected to form the I-85 user profile. During occupancy data collection, pairs of new and experienced data collectors record the same data for fifteen minute intervals of each collection period to form a comparative double-counted data set. Differences in the occupancy data are obtained from the I-85 user profile correction process to compare the

error rates between occupancy data collectors. This study is designed to identify whether new data collectors achieve the same quality of occupancy data as experienced data collectors.

A literature review is included reviewing the history, purpose, and performance measures of HOV and HOT lane systems. The methodology is provided for manual occupancy data collection, license plate video processing, occupancy and license plate data matching, and data analyses for the comparative double-count data. Results and conclusions are provided for three different occupancy sessions performed by three inexperienced and two experienced undergraduate research assistants (URAs).

Chapter 2: Literature Review

The following literature review contains information on the history, purpose, and performance measures for HOV/HOT lanes.

2.1 History of HOV Lane Systems

The first HOV lane was introduced in New Jersey in 1969 as a bus-only lane to improve transit ridership on NJ-495 [1]. At the same time, a similar bus-only lane was implemented in Virginia on I-395 to allow transit flow during a period of construction. After construction was completed, the lane was converted into two reversible HOV lanes which are still in operation today [1]. In an effort to improve air quality in Los Angeles, the general purpose (GP) left lanes were converted to a three-occupant HOV lane on the Santa Monica Freeway. After just 21 weeks, the project was cancelled due to unpopular public opinion, but a similar HOV lane was added on the El Monte Freeway as a busway facility. This HOV lane was opened to three-occupant carpools in 1976 and currently carries approximately 1,200 vehicles and 5,700

passengers per hour [1]. Since the 1970s, Seattle has opened 191 route-miles of HOV lane systems which currently carry over 100,000 commuters per day [1].

From 1993 to 2009 the number of vehicle miles travelled nationally increased by 25% while the total number of carpool and vanpool trips have decreased from almost 12 million in 1993 to roughly 10 million in 2003 [2]. Comparatively, HOV lane miles have increased from 1,300 miles in 1995 to 3,100 miles in 2005. The highest concentrations of these HOV lane miles are located in California (1000 lane miles), Georgia (400 lane miles), and Texas (300 lane miles). Although national carpooling rates have decreased, the carpooling rates in these HOV corridors have increased over 100% over the past two decades [2].

Concurrent-flow lanes are the primary type of HOV lanes used in the United States today. These lanes are adjacent to the general GP lanes and flow in the same direction, but there are indicators such as varying markers or different pavements to distinguish the HOV lane from the adjacent GP lanes. Buffer-separated concurrent-flow lanes comprise 48% of all HOV lanes and feature a physical barrier between the HOV and general purpose lanes; 28% of HOV concurrent-flow lanes are considered non-buffered [1]. The HOV/HOT lanes on I-85 in Atlanta are non-buffered concurrent-flow lanes [3].

2.2 Purpose of HOV/HOT Systems

There are currently over 130 HOV lane systems in 23 metropolitan areas of the United States [1]. These HOV lane systems are typically operational during specified peak morning and afternoon periods or on a 24-hour basis [4]. HOV lanes help mitigate congestion levels on metropolitan freeway corridors. By providing a free-flow opportunity for vehicles carrying multiple passengers, higher vehicle occupancy results on the managed lane and the number of

single-occupant vehicles on the corridor declines. Expanded areas of congestion have led to the expansion of HOV facilities to non-radial routes in the suburbs, particularly in Los Angeles, Seattle, and the San Francisco Bay area [1] .

The primary reason for installing HOV systems is the presence of high congestion levels. The general premise is that restricting HOV lanes to vehicles which meet a certain occupancy level or emissions standard will reduce the total number of vehicles on highways and subsequently increase the flow of traffic in GP lanes [4]. By reducing the number of vehicles using the corridor, HOV lanes can have the benefit of reducing travel time and the number of vehicle trips, managing travel demand, increasing carpooling rates, and improving air quality [2, 4]. An HOV corridor will induce carpooling primarily by three factors: commute time reduction, convenience of participating in a carpool, and reduced cost of carpooling relative to driving oneself [2]. The potential problem that arises on a congested corridor is when not enough vehicles use the HOV lane and single-occupant vehicles are displaced into already crowded general purpose lanes.

If an HOV lane is consistently running under capacity, an HOT lane may be implemented to allow for single or lower-occupancy-vehicles to use the lane for a fee [2]. An HOT lane can alleviate congestion during all periods of the day by adjusting usage prices to meet current traffic demands [4]. As reported by Plotz and Konduri [4], HOT lanes can be successful in efficiently allocating HOV lanes to lower-occupancy-vehicles without discouraging carpooling through tolls.

2.3 Performance Measures of HOV/HOT Systems

2.3.1 Capacity and Flow

The performance of an HOV system can be evaluated along several lines. Legislation in the United States dictates that HOV lanes must operate at a minimum of 45 mph during 90% of the peak congestion periods. If this stipulation is not met for 180 consecutive days, a policy change must be implemented [3]. The two defining characteristics of an HOV lane are its capacity and flow. A 2011 Georgia Tech study of the I-85 HOV facility in Atlanta showed that the HOV lane serves fewer vehicles per hour than the general purpose lane, but carries more passengers per hour [3]. Congestion on HOV lanes is still prevalent, but there is typically a delay in congestion compared to the general purpose lanes [5]. Single-occupant vehicle (SOV) violators, slow-moving HOV lane users, and weaving into and out of the lane also create congestion in the HOV lane [3]. The SOV violation rate is approximately 13% within the I-85 HOV corridor in Atlanta [6].

2.3.2 Occupancy

Another measure of HOV/HOT lane performance is vehicle occupancy (persons per vehicle). As previously stated, an increase in vehicle occupancy can increase the total number of I-85 users and decrease the total number of vehicle trips. Typical occupancy requirements are two or three occupants per car; free use of the HOT lane on I-85 in Atlanta currently requires three occupants. When vehicle occupancy is less than three, a fee is charged based on the current congestion conditions and lane access [3]. One concern with quantifying occupancy is the propensity for related household members to carpool in the same vehicle. This phenomenon has been termed “fampooling” and it is estimated that 43% of carpoolers fall into this category

[2, 3]. A larger survey sample of fampoolers and their use of HOV/HOT lanes must be attained to assess whether fampoolers would undergo their carpool trip regardless of an HOV/HOT lane or if increasing occupancy requirements or toll fees would have an effect on fampool numbers [2].

The purpose of this analysis is to assess the quality of vehicle profile data collected by URAs on I-85 in Atlanta with emphasis on the left GP lane. The analysis aims to assess the accuracy of the vehicle flows, vehicle order, and vehicle classifications collected by comparing the entries to simultaneous highway videos, license plate values, and vehicle registration databases. The license plate values are matched with the registration database and are subsequently verified using the corresponding highway videos. This corrected data set is the control for the analysis used in assessing the quality of URA data collection.

Chapter 3: Methodology

3.1 Data Collection Methods

The data collection process occurs in the morning peak period (7:00am to 9:00am) and in the afternoon peak period (4:30pm to 6:30pm). Data are collected quarterly and only the peak traffic direction is observed; i.e. during the morning shift southbound traffic is observed and during the afternoon shift northbound traffic is observed.

The data collection sites are Chamblee Tucker Road (CTR), Jimmy Carter Boulevard (JCB), Beaver Run Road (BRR), Pleasant Hill Road (PHR), and Old Peachtree Road (OPR). The equipment used in the process is a set of tripods for the cameras, high definition video-cameras with high capacity SD memory cards, cable tethers and locks for the cameras, netbook laptops, customized keypads for use with the netbooks, safety vests, a mallet and stake for

supporting the tripods, and storage containers for all the equipment. The students wear the safety vests during all points of the data collection process on the highway. On certain occasions contractible lawn chairs will be used by the students for comfort.

3.1.1 License Plate Data Collection - Apparatus Configuration

On the bridge crossing the highway at each site, the graduate students and possibly several undergraduate students will expand and place one tripod between each lane, one tripod on each end of the outside lanes, and one central tripod to view all the lanes. Figure 1 below displays the camera setup on the bridge with the nearest camera being the outside lane, followed by the interior lane and central cameras, and the far right lane camera. This configuration may change due to a demand for certain data or a limit in the number of cameras available.



Figure 1: Example of configurations for the bridge cameras.
Photo Credit: Katie Smith

The cameras are placed on the tripods and adjusted for each specific position. Cameras on the outside lanes are adjusted so the edge of the video nearest to the other lanes will be the

dotted white lines of the outside lane. The other edge of the video will be the solid line indicating the edge of the highway. This configuration ensures that outside lanes such as the HOV/HOT and far right lanes are displayed individually. The interior cameras are adjusted to display two lanes by centering the dotted white line dividing each lane. The interior and outside cameras are vertically adjusted and zoomed in a manner such that a view down the highway with the flow of traffic is attained (as shown in Figure 2 below). The central camera is placed in the center of the highway and adjusted so the video displays every lane and is completely zoomed out.



Figure 2: Example of camera setup and configuration with highway view.
Photo Credit: Katie Smith

Once each camera is set up and adjusted, the cable tether and lock are used for safety to restrict the movement of the tripods from weather conditions, shaking of the bridge, human interference, or any other factor that might shake or knock over the tripod. Figure 2 above displays the camera and cable tether setup with a view of the highway. After a final check of the camera and tripod configurations, the student presses the record button and collects two hours of

video data. The cameras are checked for proper function and configuration periodically over the two hours and are shut down after two hours of recording. When recording is finished the camera, tripods, tethers, and locks are disassembled and stored in their proper containers.

3.1.2 Occupancy Data Collection

While the students on the bridge set up the cameras, the remaining undergraduate students cross the street and find a suitable area on the grass banks next to the highway for the process of manual data collection. The students sit on the opposite side of the bridge from the cameras so that manual data are collected before the bridge and video data are collected after the cars pass under the bridge. A tripod is set up and a gore area camera is attached to capture video of the entire highway before the bridge. Gore area cameras are set up in the same manner as the central camera on the bridge, except the video is taken roughly at an angle 30° counterclockwise from a perpendicular view of highway flow. This tripod must be tethered to the ground using the stake and mallet. The stake is hammered into the ground using the mallet and is placed next to the tripod. Once in the ground, the cable and lock are used to tether the tripod to the stake to prevent movement of the camera and tripod. Figure 3 below displays this gore area camera configuration and provides a side view of the highway. The camera is adjusted to display all the highway lanes and zoomed out to provide a view looking up the highway opposite to traffic flow. When completely adjusted, a student begins recording and continues for two hours while periodically checking for proper camera configuration until the two hours are up. After two hours of recording, the camera, tripod, cable tether, lock, and stake are disassembled and stored in their proper containers.



Figure 3: Example of gore area camera and stake setup.
Photo Credit: Katie Smith

Once the camera is set up and a suitable vantage point on the grass bank is found, each student is assigned to a lane to collect occupancy data. Depending on the deployment plan for a specific site, an extra student may be assigned to the HOT lane or to other lanes. The field supervisor uses the clipboard data sheet to record the time of collection, vehicle type, number of passengers, and any notes on the first and last cars of the target lane with the extra student.

Each student is provided a netbook with a number corresponding to their lane and a keypad for entering data. The keypad, shown in Figure 4, contains values for the type of vehicle, number of passengers in the vehicle, and several data entry buttons. Vehicle types are Light Duty Vehicle (LDV), Sport Utility Vehicle (SUV), or Heavy Duty Vehicle (HDV); these vehicle types are classified using the body types listed in Table 1.0. The options for number of passengers are 1, 2, or 3, indicating the student is sure of the number, and 1+, 2+, 3+, 4+, indicating the student is anything short of positive about the number of passengers. A “Miss”

key is included in case the student completely misses a vehicle and a “Clear” key will remove the last entry from the data. The “Enter” key will input the data into the computer.

Table 1.0: Vehicle Body Type Re-Classification
Table Credit: Katie Smith

LDV	SUV	HDV
2S (2 door sedan)	AM (ambulance)	HR (horse trailer)
3S (3 door sedan)	CT (camper trailer)	TL (trailer)
4S (4 door sedan)	MP (multi-purpose)	UL (trailer)
5S (5 door sedan)	TK (pick-up truck)	BU (bus)
CN (convertible)	TR (pick-up truck)	
CP (coupe)	VN (van)	
LM (limousine)	WK (work truck)	
MC (motorcycle)	JP (jeep)	
RD (roadster)	BT (boat trailer)	

		C	
1	1+	HDV	SUV
2	2+		LDV
3	3+		M
4+			I
			S
			S

Figure 4: Keypad for Occupancy Data Collection Process.
Photo Credit: Katie Smith

The students start the netbooks and open the “Collect Occupancy” folder on the Desktop. From here they launch the program and enter their name and the file name in the format SITE_VO_AM/PM_NB/SB_LANE#_MMDDYY. Once the program has started the student

finds his or her lane and begins collecting data by recording first the vehicle classification, then the number of passengers, then pressing the enter button. The student will continue this process for every car that they can observe until two hours have passed, at which time the data collection stops. If a student misses a car or makes a mistake the Miss and Clear buttons are used and data collection continues as normal. When the two hours are up, the student presses “Q” to quit the program, shuts down the netbook, and returns the netbooks and keypads to their proper containers. The students gather all the equipment and return to their vans or cars.

3.1.3 Double-count Occupancy Data Collection

A comparative set of occupancy data are sometimes created using pairs of one experienced and one inexperienced data collector, or two experienced data collectors. For fifteen minute intervals of each two-hour collection period, one experienced and one inexperienced or new data collector monitor the same lane simultaneously. Each collector follows the methods detailed in Section 3.1.2 to obtain occupancy data. This results in a set of double-counted data to be analyzed using methods described later in this report.

3.1.4 License Plate Data Processing

The video processing component of the data collection consists of recording the license plate for every vehicle from the video data recorded on the bridge. Once the SD cards are removed from the cameras after data collection is finished, files are uploaded to the laboratory computers and then converted from continuous video into a picture slideshow format. After opening the program in the computer lab and entering his or her name, the student will be looking at a freeze-frame picture of the video data for a certain lane or two lanes. A sample of

the video processing program with a view of the HOT lane and the far left GP lane is shown below in Figure 5.

The video can be advanced or reversed using the arrow keys. The vehicle information categories in Figure 5 from top to bottom are the license plate value, vehicle type, license plate state, and a comments section. Data processors fill out all information to the best of their abilities and hit the enter button when finished. If a license plate is not discernible, the student will put “M” for miss in the license plate section and “Missed” in the state section while providing a comment explaining why the plate was missed.

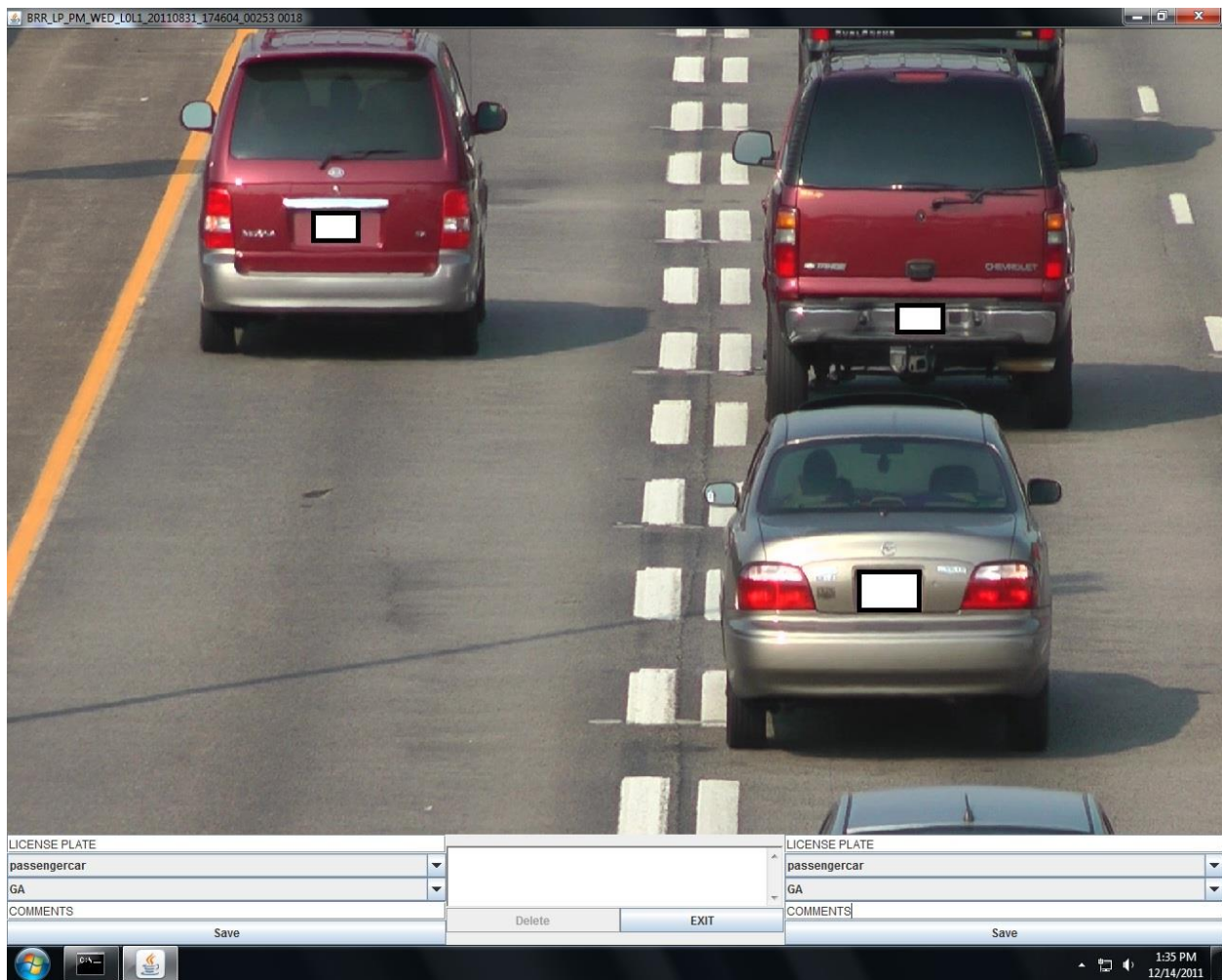


Figure 5: Sample screenshot of video processing program with a view of the HOT lane and adjacent lane.

3.2 Matching Occupancy Data with License Plate Data

The following methodology is explained in further detail in a recent M.S. thesis by Smith entitled “A Profile of HOV Lane Vehicle Characteristics on I-85 Prior to HOV-to-HOT Conversion” [3]. To match the occupancy and license plate data, three common variables are utilized: lane number, vehicle classification, and time gap between vehicles. After the license plate data are processed using the methods previously described in Section 3.1.4, the vehicle classification, make, and model for each vehicle are obtained from the vehicle registration database. The license plate video is then processed a second time using the information from the vehicle registration database to verify the accuracy of the original license plate data. On average, 11% of license plate data were corrected to account for vehicle omissions, license plate misspellings, or other errors. The gore area camera, used during the occupancy data collection described in Section 3.1.2, is monitored simultaneously with the license plate video as a means of accurately maintaining the order of vehicles. The order of vehicles can be affected by vehicle lane changes, visibility concerns of occupancy data collectors, or other factors to be discussed later in this report [3].

Once occupancy and license plate data are reviewed and errors are removed, the three data sets are matched using the shared variables of vehicle classification and time gap between vehicles. As observed by Smith [3], the occupancy and license plate data sets have different time gaps for each record due to differences in the camera and netbook clocks or varying distances between collection points; these time gaps are relatively consistent considering the furthest distance between collection points is roughly one-third of a mile. As previously stated in Section 3.1.2 and Section 3.1.4, vehicle classification is recorded for both the occupancy and license

plate data. Vehicle classification is used with comments from the re-processing of the two video data sets to match the gore area occupancy video, manual occupancy, and license plate video data into one comprehensive profile of I-85 users. Figure 6 below depicts the matching of the three data sets in a flow chart [3].

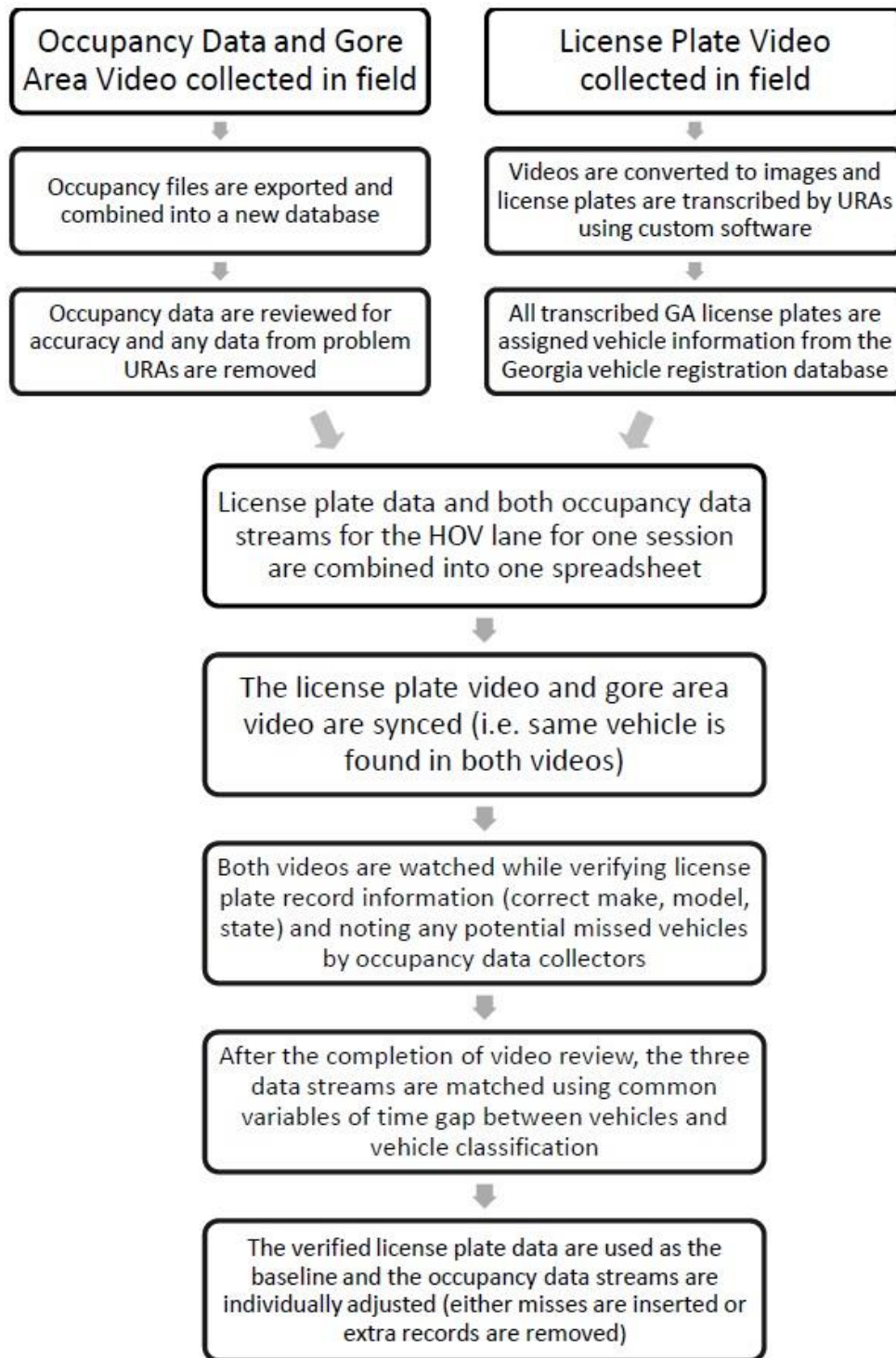


Figure 6: Flow Chart of Matching Occupancy and License Plate Data.
Chart Credit: Katie Smith

Figure 7 is a sample screenshot of all three data streams matched and corrected for errors, omissions, or other mistakes. The left columns represent the URA (URA AX), the center columns the URA monitor (URA BX), and the right columns the corrected license plate data. Vehicle classification is used as the primary tool in matching with the license plate data serving as the veritable order and identity of vehicles. As previously mentioned the three data streams vary in timestamps; therefore this variable is used mainly as a secondary tool for matching the data streams.

Once the data are organized into the same spreadsheet, the start and end times are matched for the monitor and the license plate data. Although there is no correct or incorrect method of matching these streams, the typical methodology used in this thesis is to find: (1) two or more consecutive HDV vehicles, (2) more than four consecutive LDV or SUV vehicles, or (3) any other distinct combination to be used as a baseline for the matching process. When one of these conditions is met, the time gaps are checked to confirm the match. As seen in Figure 7, the time gap columns are given a 3-color percentile scale from green (long gap) to yellow (medium gap) to red (short or no gap). If the time gaps are consistent for the selection of matched vehicle classifications, that selection is said to be completely matched. Subsequent matches are based off the selection and every vehicle missed, switched in vehicle order, incorrectly classified, or irregularly added is noted and the mistake is corrected. When the license plate data and monitor data are matched from start to finish, the same process is repeated with the URA data and the license plate data. The amount of error is calculated for both the URA and the monitor to identify the reliability of their data.

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	
847	MISS					LDV							1.33E+09	Andrew L	7:33:21 AM	00:00.9	LDV		1		7:33:31 AM	00:01.0	LDV				45	20120207	
848		1.33E+09	AKSHI	7:34:07 AM	00:01.7	LDV							1.33E+09	Andrew L	7:33:22 AM	00:01.5	LDV		1		7:33:32 AM	00:01.0	LDV				45	20120207	
849		1.33E+09	AKSHI	7:34:09 AM	00:04.3	HDV							1.33E+09	Andrew L	7:33:26 AM	00:04.3	HDV		1		7:33:35 AM	00:04.0	HDV				UL	20120207	
850		1.33E+09	AKSHI	7:34:13 AM	00:05.2	SUV							1.33E+09	Andrew L	7:33:30 AM	00:05.2	SUV		1		7:33:39 AM	00:05.0	SUV				TK	20120207	
851		1.33E+09	AKSHI	7:34:18 AM	00:00.9	HDV							1.33E+09	Andrew L	7:33:35 AM	00:00.9	HDV		1		7:33:44 AM	00:01.0	HDV				UL	20120207	
852		1.33E+09	AKSHI	7:34:19 AM	00:01.7	LDV							1.33E+09	Andrew L	7:33:36 AM	00:00.9	LDV		1		7:33:45 AM	00:01.0	LDV				45	20120207	
853		1.33E+09	AKSHI	7:34:21 AM	00:01.7	LDV							1.33E+09	Andrew L	7:33:37 AM	00:01.5	LDV		1		7:33:46 AM	00:04.0	LDV				45	20120207	
854		1.33E+09	AKSHI	7:34:23 AM	00:00.0	LDV							1.33E+09	Andrew L	7:33:40 AM	00:00.8	LDV		1		7:33:50 AM	00:01.0	949RLV	LDV	TOYT	COROLLA	2004	45	20120207
855		1.33E+09	AKSHI	7:34:23 AM	00:13.0	LDV							1.33E+09	Andrew L	7:33:41 AM	00:12.1	LDV		1		7:33:51 AM	00:10.0	LDV				45	20120207	
856		1.33E+09	AKSHI	7:34:36 AM	00:02.6	LDV							1.33E+09	Andrew L	7:33:53 AM	00:01.7	LDV		1		7:34:01 AM	00:03.0	LDV				45	20120207	
857		1.33E+09	AKSHI	7:34:38 AM	00:00.9	LDV							1.33E+09	Andrew L	7:33:55 AM	00:00.9	LDV		1	COP, LP 0020	7:34:04 AM	00:01.0	LDV				45	20120207	
858		1.33E+09	AKSHI	7:34:39 AM	00:00.9	SUV							1.33E+09	Andrew L	7:33:56 AM	00:00.9	SUV		1		7:34:05 AM	00:01.0	SUV				TK	20120207	
859		1.33E+09	AKSHI	7:34:40 AM	00:05.2	SUV							1.33E+09	Andrew L	7:33:57 AM	00:05.2	SUV		1		7:34:06 AM	00:05.0	SUV				TK	20120207	
860		1.33E+09	AKSHI	7:34:45 AM	00:11.2	SUV							1.33E+09	Andrew L	7:34:02 AM	00:11.2	SUV		1		7:34:11 AM	00:11.0	SUV				TK	20120207	
861		1.33E+09	AKSHI	7:34:56 AM	00:04.3	LDV							1.33E+09	Andrew L	7:34:13 AM	00:00.9	LDV		1		7:34:22 AM	00:01.0	LDV				45	20120207	
862	MISS					SUV							1.33E+09	Andrew L	7:34:14 AM	00:06.0	SUV		1		7:34:23 AM	00:06.0	SUV				MP	20120207	
863		1.33E+09	AKSHI	7:35:01 AM	00:05.2	HDV							1.33E+09	Andrew L	7:34:20 AM	00:02.6	HDV		1		7:34:29 AM	00:04.0	HDV				UL	20120207	
864	MISS		1.33E+09	miss	7:35:06 AM	00:00.9	SUV						1.33E+09	Andrew L	7:34:23 AM	00:01.7	SUV		1		7:34:33 AM	00:00.0	AGA8426	SUV			MP	20120207	
865	MISS		1.33E+09	AKSHI	7:35:07 AM	00:00.9	SUV						1.33E+09	Andrew L	7:34:24 AM	00:00.9	SUV		1		7:34:33 AM	00:01.0	SUV				MP	20120207	
866		1.33E+09	AKSHI	7:35:08 AM	00:01.7	LDV							1.33E+09	Andrew L	7:34:25 AM	00:00.9	LDV		1		7:34:34 AM	00:01.0	LDV				45	20120207	
867		1.33E+09	AKSHI	7:35:09 AM	00:06.9	SUV							1.33E+09	Andrew L	7:34:26 AM	00:05.2	SUV		1		7:34:35 AM	00:06.0	LDV				45	20120207	
868		1.33E+09	AKSHI	7:35:16 AM	00:01.7	HDV							1.33E+09	Andrew L	7:34:31 AM	00:03.5	HDV		1		7:34:41 AM	00:04.0	HDV				UL	20120207	
869		1.33E+09	AKSHI	7:35:18 AM	00:08.6	HDV							1.33E+09	Andrew L	7:34:35 AM	00:03.5	HDV		1		7:34:45 AM	00:02.0	HDV				UL	20120207	
870	WRONG		1.33E+09	AKSHI	7:35:27 AM	00:01.7	SUV						1.33E+09	Andrew L	7:34:38 AM	00:06.0	LDV		2		7:34:47 AM	00:06.0	LDV				45	20120207	
871	WRONG		1.33E+09	AKSHI	7:35:28 AM	00:00.9	SUV						1.33E+09	Andrew L	7:34:44 AM	00:01.7	SUV		1		7:34:53 AM	00:01.0	SUV				TK	20120207	
872	WRONG		1.33E+09	AKSHI	7:35:29 AM	00:00.9	SUV						1.33E+09	Andrew L	7:34:44 AM	00:01.7	SUV		1		7:34:54 AM	00:01.0	LDV				45	20120207	
873	MISS		1.33E+09	AKSHI	7:35:30 AM	00:04.3	SUV				MISS			1.33E+09	Andrew L	7:34:46 AM	00:06.0	SUV		1		7:34:55 AM	00:07.0	SUV				VN	20120207
874		1.33E+09	AKSHI	7:35:34 AM	00:01.7	SUV							1.33E+09	Andrew L	7:34:52 AM	00:00.9	SUV		1		7:35:02 AM	00:00.0	5741AHW	SUV	CHEV	K15	2002	MP	20120207
875		1.33E+09	AKSHI	7:35:36 AM	00:00.9	LDV							1.33E+09	Andrew L	7:34:53 AM	00:00.9	LDV		1		7:35:02 AM	00:02.0	77008E	LDV	HOND	CIVIC U.S.	1998	CP	20120207
876		1.33E+09	AKSHI	7:35:37 AM	00:00.9	LDV							1.33E+09	Andrew L	7:34:54 AM	00:00.9	LDV		1		7:35:04 AM	00:01.0	LDV				45	20120207	
877		1.33E+09	AKSHI	7:35:38 AM	00:03.5	LDV							1.33E+09	Andrew L	7:34:55 AM	00:04.3	LDV		1		7:35:05 AM	00:04.0	AGT0861	LDV	TOYT	CAMRY U.	1999	45	20120207
878		1.33E+09	AKSHI	7:35:41 AM	00:04.3	LDV							1.33E+09	Andrew L	7:34:59 AM	00:05.2	LDV		1		7:35:09 AM	00:04.0	ADR2951	LDV	LINC	TOWN CA	1992	45	20120207
879		1.33E+09	AKSHI	7:35:46 AM	00:03.5	LDV							1.33E+09	Andrew L	7:35:04 AM	00:01.7	LDV		2		7:35:13 AM	00:03.0	LDV				45	20120207	
880		1.33E+09	AKSHI	7:35:49 AM	00:04.3	LDV							1.33E+09	Andrew L	7:35:06 AM	00:04.3	LDV		1		7:35:16 AM	00:01.0	AGA3497	LDV			45	20120207	
881		1.33E+09	AKSHI	7:35:53 AM	00:00.9	LDV							1.33E+09	Andrew L	7:35:10 AM	00:00.9	LDV		1		7:35:19 AM	00:01.0	LDV				45	20120207	
882		1.33E+09	AKSHI	7:35:54 AM	00:06.0	LDV							1.33E+09	Andrew L	7:35:11 AM	00:05.2	LDV		1		7:35:20 AM	00:06.0	LDV				45	20120207	
883		1.33E+09	AKSHI	7:36:00 AM	00:00.9	LDV							1.33E+09	Andrew L	7:35:16 AM	00:00.9	LDV		1	Switched 544	7:35:26 AM	00:00.0	LDV				45	20120207	
884		1.33E+09	AKSHI	7:36:01 AM	00:02.6	SUV							1.33E+09	Andrew L	7:35:17 AM	00:01.7	SUV		1		7:35:26 AM	00:02.0	SUV				TK	20120207	
885		1.33E+09	AKSHI	7:36:04 AM	00:02.6	LDV							1.33E+09	Andrew L	7:35:19 AM	00:03.5	LDV		1		7:35:28 AM	00:04.0	LDV				45	20120207	
886		1.33E+09	AKSHI	7:36:06 AM	00:00.0	LDV							1.33E+09	Andrew L	7:35:22 AM	00:00.9	LDV		1		7:35:32 AM	00:01.0	1457AAE	LDV	MTS	GALANT D	2006	45	20120207
887																			WRONG, DUF	7:35:33 AM	00:00.0	BY1087	SUV				VN	20120207	
888		1.33E+09	AKSHI	7:36:06 AM	00:02.6	SUV							1.33E+09	Andrew L	7:35:23 AM	00:06.0	SUV		1		7:35:33 AM	00:05.0	BY1087	SUV	FORD	EXPEDITIC	2002	MP	20120207
889		1.33E+09	AKSHI	7:36:09 AM	00:05.2	HDV							1.33E+09	Andrew L	7:35:29 AM	00:01.7	HDV		1		7:35:38 AM	00:05.0	BY1087	SUV				20120207	
890		1.33E+09	AKSHI	7:36:14 AM	00:01.7	SUV							1.33E+09	Andrew L	7:35:31 AM	00:00.9	SUV		1		7:35:40 AM	00:01.0	LDV				MP	20120207	
891		1.33E+09	AKSHI	7:36:16 AM	00:00.9	SUV							1.33E+09	Andrew L	7:35:32 AM	00:02.6	SUV		1		7:35:41 AM	00:03.0	AFAS689	SUV	TOYT	HIGHLAND	2004	MP	20120207
892		1.33E+09	AKSHI	7:36:17 AM	00:05.2	HDV							1.33E+09	Andrew L	7:35:34 AM	00:04.3	HDV		1		7:35:44 AM	00:05.0	HDV				UL	20120207	
893		1.33E+09	AKSHI	7:36:27 AM	00:00.0	LDV							1.33E+09	Andrew L	7:35:39 AM	00:01.3	HDV		1		7:35:49 AM	00:01.0	SUV				MP	20120207	

Figure 7: Sample screenshot of all three data streams matched and corrected

Chapter 4: Results

Upon completion of the matching process, the number of missed, extra, switched, or incorrectly classified vehicles is tallied for the two sets of occupancy data. A missed vehicle is one which appears in the license plate and/or gore area videos that is not accounted for in the occupancy data. Switched vehicles are two or more vehicles that are incorrectly ordered in the occupancy data; e.g. license plate video shows one HDV followed by two SUVs but the occupancy data for URA A1 shows one SUV, followed by the HDV, followed by the final SUV. An extra vehicle is one which appears in the occupancy data but not in the license plate and/or gore area videos. Incorrectly classified vehicles in the occupancy data are those incorrectly classified according to Table 1.0.

A summary of the session information, start/end time, sample duration, total entries, and the classification of entries is listed in Table 2.0 for each session from this analysis; a “Miss”

value is proper use of the miss-button while a “Wrong/Miss” is improper or excessive use of the miss-button. Table 2.0 summarizes the field data used in the analysis and provides the raw entries for each URA, date, and time. Table 2.1 displays the analysis results for each URA and defines each entry type as either perfectly matched, missed vehicle, extra vehicle, or a vehicle classification error (i.e. light duty vehicle recorded as sport utility vehicle or vice versa). Missed vehicles are counted for each URA during the analysis data comparison to assure analysis vehicle totals are consistent with actual vehicle totals. Table 2.1 also introduces the effective entry totals considered in the analysis by eliminating the missed and extra vehicles for each URA. Missed and extra vehicles indicate a lack of vehicle match but are not considered errors. Table 2.2 displays the distribution of entry types as a percentage of the total perfect matches, extra vehicles, and vehicle classification errors. Table 2.3 eliminates the extra vehicles and displays the distribution of effective entries. The rate of misclassification is considered the error rate for each URA. Figure 8 displays the total effective entries for each URA featured in the analysis. Figure 9 shows the error rates by type for each URA and monitor.

The overall error rate percentage for the sample (i.e. from all URAs and monitors combined) was 4.9% from a total of 2089 effective occupancy data entries. On average, the URAs produced error rates 2.4 percentage points higher than the monitors. Among the URAs, URA A2 produced the lowest error rate of 1.8% and URA A3 produced the highest error rate of 3.3%. Among the monitors, URA B1 had an average error rate of 5.4% while URA B2 produced the lowest error rate of 0.3%. Among the classification errors, SUVs misclassified as LDVs were the most frequent error at 2.6% of the total effective entries followed by LDVs misclassified as SUVs at 1.8% of the total effective entries. Misclassifying HDVs as SUVs occurred far less frequently at 0.4% of the total effective entries.

Table 2.0: Matched License Plate Videos and Occupancy Data

Session	<i>BRR - 2/23/12 - AM - Lane 1</i>			<i>OPR - 2/7/12 - AM - Lane 3</i>			<i>CTR - 1/19/12 - PM - Lane 1</i>		
Data Source	LP	URA A1	URA B1	LP	URA A2	URA B2	LP	URA A3	URA B1
Start	7:47:31 AM	7:45:34 AM	6:46:42 AM	7:28:03 AM	7:21:42 AM	7:20:58 AM	5:05:01 PM	5:10:39 PM	5:17:13 PM
End	8:01:09 AM	7:59:13 AM	7:00:21 AM	7:36:59 AM	7:37:32 AM	7:36:50 AM	5:23:32 PM	5:29:14 PM	5:35:38 PM
Duration	0:13:38	0:13:39	0:13:39	0:08:56	0:15:50	0:15:52	0:18:31	0:18:35	0:18:25
Total Entries	453	405	408	310	287	308	388	345	379
LDV	247	225	220	144	124	142	192	177	191
SUV	205	179	184	110	97	107	194	137	183
HDV	1	1	2	56	56	57	2	3	4
Miss	N/A	0	2	N/A	7	2	N/A	24	1
Wrong/Miss	N/A	0	0	N/A	3	0	N/A	4	0

Table 2.1: Analysis Data – Entry Totals by Type

Entry Type Distribution & Analysis Vehicle Totals								
Entry Type		Symbol	<i>BRR - 2/23/12 - AM - Lane 1</i>		<i>OPR - 2/7/12 - AM - Lane 3</i>		<i>CTR - 1/19/12 - PM - Lane 1</i>	
			URA A1	URA B1	URA A2	URA B2	URA A3	URA B1
Perfect Match		P	382	392	278	307	292	335
Missed Vehicle		M	55	53	26	2	56	21
Extra Vehicle		E	7	8	3	0	13	12
Vehicle Classification Error	<i>SUV for LDV</i>	C1	6	4	3	0	13	12
	<i>LDV for SUV</i>	C2	10	3	2	0	26	19
	<i>HDV for SUV</i>	C3	0	1	1	1	1	1
Analysis Total Entries = P + E + $\sum C$		Ta	405	408	287	308	345	379
Actual Total Entries		T	405	408	287	308	345	379
Effective Total Entries = Ta - E		Te	398	400	284	308	332	367
Analysis Vehicle Total = Te + M		Aa	453	453	310	310	388	388
Actual Vehicle Total		A	453		310		388	

Table 2.2: Analysis Data – Entry Percentages by Type

Analysis Total (Ta) Entry Type Percentages									
Entry Type (% of Ta)		Symbol	Site - Date - Session - Lane	BRR - 2/23/12 - AM - Lane 1		OPR - 2/7/12 - AM - Lane 3		CTR - 1/19/12 - PM - Lane 1	
			URA XX	URA A1	URA B1	URA A2	URA B2	URA A3	URA B1
			Ta	405	408	287	308	345	379
Perfect Match		P		94.32	96.08	96.86	99.68	84.64	88.39
Extra Vehicle		E		1.73	1.96	1.05	0.00	3.77	3.17
Vehicle Classification Error	SUV for LDV	C1		1.48	0.98	1.05	0.00	3.77	3.17
	LDV for SUV	C2		2.47	0.74	0.70	0.00	7.54	5.01
	HDV for SUV	C3		0.00	0.22	0.32	0.32	0.26	0.26
	Σ C		3.95	1.94	2.06	0.32	11.56	8.44	
Sum = P + E + Σ C				100	100	100	100	100	100

Table 2.3: Analysis Data – Effective Entries by Type

Effective Total (Te) Entry Type Percentages									
Entry Type (% of Te)		Symbol	Site - Date - Session - Lane	BRR - 2/23/12 - AM - Lane 1		OPR - 2/7/12 - AM - Lane 3		CTR - 1/19/12 - PM - Lane 1	
			URA XX	URA A1	URA B1	URA A2	URA B2	URA A3	URA B1
			Te	398	400	284	308	332	367
Perfect Match		P		95.98	98.00	97.89	99.68	87.95	91.28
Vehicle Classification Error	SUV for LDV	C1		1.51	1.00	1.06	0.00	3.92	3.27
	LDV for SUV	C2		2.51	0.75	0.70	0.00	7.83	5.18
	HDV for SUV	C3		0.00	0.25	0.35	0.32	0.30	0.27
	Σ C		4.02	2.00	2.11	0.32	12.05	8.72	
Sum = P + Σ C				100	100	100	100	100	100

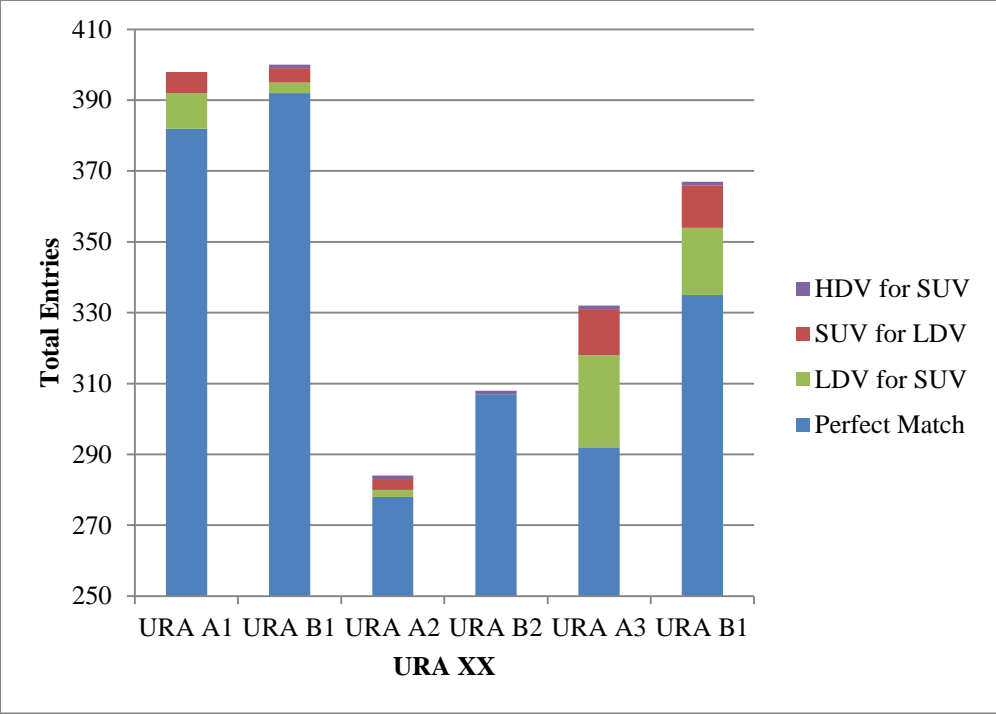


Figure 8: Total Effective Entries by Type

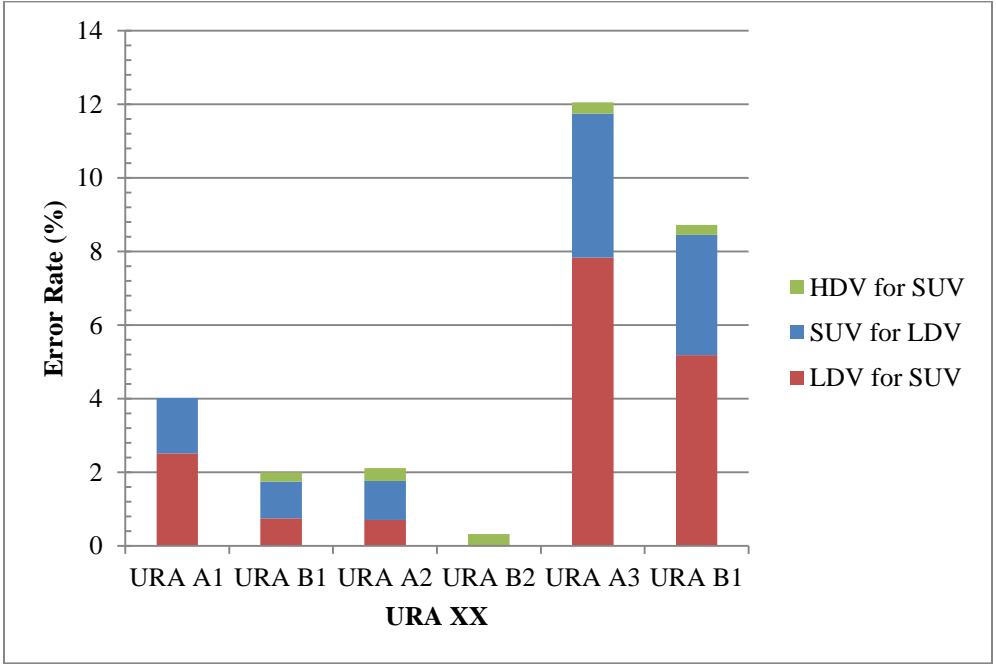


Figure 9: Total Error Rates by Type

Chapter 5: Discussion and Conclusion

5.1 Discussion of Results

Due to the nature of occupancy data collection, missed and extra vehicles are expected and therefore are ignored in the analysis. An overall error rate below 5% for the entire sample shows that URAs are capably performing their tasks but there is room for improvement to ensure LDVs and SUVs are properly classified as frequently as possible. As expected, the monitors produced less error than the new URAs but the new URAs were all within four percentage points of their respective monitors. Further analysis of the URAs over extended periods is recommended to assess if the gap between new and experienced URAs diminishes with further data collection and/or a refresher-course on the proper vehicle classifications. The primary concern arising from this analysis is the discrepancy between the error rates from BRR and OPR in the morning session and those from CTR in the afternoon session. The average error rate for the afternoon session was over ten percentage points higher than the average from the two morning sessions. The sample size of this analysis is too small to establish a relevant relationship between the morning and afternoon sessions or between the sites. It is recommended that this analysis be repeated to include a larger sample of morning and afternoon sessions across all collection sites to analyze whether the time of day, amount of sunlight, site access restrictions, sight restrictions, or any other factor is consistently affecting the quality of data collection.

The high number of extra vehicles added by URA B1 and URAs A1 & A3 could have been caused by over-estimating the number of cars in a long succession or by guessing vehicles when site restrictions or concentration lapses occurred. Excluding URA B2, all the URAs

frequently misclassified LDVs as SUVs and vice versa. This trend suggests a discrepancy between what the URAs define as LDVs & SUVs and what Table 1.0 defines as LDVs & SUVs.

5.2 Concluding Remarks

As mentioned in the discussion, the rate of vehicle misclassification among both URAs and monitors is reason to retrain URAs on the classification system for occupancy data collection; this retraining would be simple and relatively cost-free while reducing the errors in misclassification. Although misclassification occurs roughly every 20th vehicle, whether mixing up LDVs and SUVs has a significant impact on daily highway occupancy and passenger throughput cannot be assessed by this analysis. A relationship between vehicle class and occupancy must be established through further data collection analyses to establish if vehicles with high occupancy capacities like SUVs consistently transport more passengers than LDVs on I-85. The resources available for this analysis were not adequate to establish a control for the occupancy of each vehicle. An additional gore area or bridge camera could be used to zoom-in on the far left GP lane in order to estimate the occupancy in each vehicle. Users of the HOV system could also be required to register their expected total occupancy when they register to use the lane in order to establish an average occupancy for the typical I-85 user profile.

Due to the high speed of vehicles, site restrictions, and the limits of human data collection, there is little that can be actively done to mitigate problems arising from extra, switched, or missed vehicles. A possible solution to this problem is to establish one or more additional data streams of the highway such as high-definition video cameras permanently on-site (e.g. attached to a light pole) or stationing URAs at both the side of the highway and overlooking the highway from the bridge. Live monitoring of the URAs to ensure they are

focused and attentive would reduce these types of errors but is extremely difficult if not impossible. A solution is to provide a reward system to URAs for producing low rates of misclassification over extended periods. If more error analyses like this thesis are performed in a timely fashion (relative to when occupancy data are collected), URAs can be notified of their performance and improvements can be made in classification abilities and attentiveness.

As mentioned in the discussion, the high number of extra vehicles suggests many URAs are guessing vehicles for any number of reasons. If a large trailer blocks their view or they suddenly have to sneeze it is best that they simply do not enter a vehicle as opposed to randomly inputting a non-existent vehicle. Although the URAs may feel entering a wrongly classified vehicle is better than no vehicle at all, it is difficult to perform an analysis such as this thesis when random entries appear in the data with no explanation. This problem could possibly be resolved by instilling in the URAs a desire to avoid guessing altogether. The license plate slideshow and gore area videos will show whether there existed a sight restriction when there seems to be a large string of missed vehicles. If the URAs never input extra vehicles then there are only perfect matches, misclassifications, and missed vehicles in the data set. Providing additional training to all URAs designed to ensure that they never input extra vehicles should simplify the data matching process and eliminate the need of the analyst to guess when the URAs themselves are guessing.

This report was limited to only three occupancy data collection sessions, but further analyses are recommended for the remaining occupancy data sets. The sample size should be increased to include more sites, more PM-sessions, and more URAs to account for the variability in occupancy data caused by site access, site restrictions due to time of day or weather, site restrictions due to distance from observed lanes, competence of URAs, or other factors.

The collection of I-85 User Profile data itself could also be significantly improved to help synchronize the license plate videos and data, the gore area videos, and the two streams of occupancy data. Due to the differences between the clocks on the netbooks used in occupancy and license plate collection, none of the time stamps could be used as a reference point. This caused excessive time to be spent trying to identify when the monitors began their fifteen-minute collection and rendered the monitors' manual start/end times useless. A possible way to alleviate this problem would be to train all URAs to collect data only on a small strip of each lane.

Another problem arose from the unreliability of the license plate data. Although the data were considered "corrected" and had been reviewed by a Graduate Research Assistant (GRA), there were few data sets with accurate and thorough license plate data. The three matched sessions were only successfully combined after hours of license plate data were manually checked and corrected. Matching between the license plate and occupancy data was carried out using vehicle classification and order, but if the license plate values could not be discerned then the processors simply moved on to the next car; this practice makes the license plate data almost unusable (without correction) in terms of matching abilities. If the license plate data processors are told to record as much information as possible for every entry, this problem might not arise. The entire license plate collection process could be avoided in the future with the use of automated license plate reading technology.

The primary limitation in this analysis stems from the subjectivity involved in assessing when, with what mechanism, and how frequently errors propagate in the occupancy data. Once the license plate data are corrected the baseline is established for comparison with the two occupancy data sets. Although it is clear when there are discrepancies between the baseline and the occupancy data, it is not clear exactly which entry is incorrect and whether the entry was

added, missed, misclassified, or switched. It is the judgment of the person performing the analysis which determines the frequency and mechanisms of error. It is impossible to know what the URA was thinking and doing when the mistake occurred and therefore the analyzer must simply recognize that a problem exists and determine the mechanism of error through comparison. Comparison is limited to the entries before/after the entry in question, to the other occupancy data set, and to the baseline license plate data. For these reasons the overall number of errors will be reasonably calculated but the distribution of error types will be mostly subjective.

The overall consensus from the three sessions analyzed is that URAs considered to be experienced and inexperienced alike are not exhibiting significant differences in misclassifying vehicles but all URAs appear on occasion to be unnecessarily guessing vehicles. It is recommended that this analysis be continued to include the remaining data sets from the winter 2012 data collections to verify the significance of the results presented in this analysis. Implementation of a re-training program and continuous monitoring of URAs will also assist in improving the accuracy of occupancy data and consequently the I-85 User Profile.

Bibliography

- [1] C. Fuhs and J. Obenberger, "Development of High-Occupancy Vehicle Facilities: Review of National Trends," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1781, pp. 1-9, 2002.
- [2] M. W. Burris, D. H. Ungemah, M. Mahlawat, and M. S. Pannu, "Investigating the Impact of Tolls on High-Occupancy-Vehicle Lanes Using Managed Lanes," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2099, pp. 113-122, 2009.
- [3] K. S. Smith, "A Profile of HOV Lane Vehicle Characteristics on I-85 Prior to HOV-to-HOT Conversion," Master of Science in the School of Civil and Environmental Engineering, School of Civil & Environmental Engineering, Georgia Institute of Technology, 2011.
- [4] J. Plotz, K. C. Konduri, and R. M. Pendyala, "To What Extent Can High-Occupancy Vehicle Lanes Reduce Vehicle Trips and Congestion?," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, pp. 170-176, 2010.
- [5] A. Guin, M. Hunter, and R. Guensler, "Analysis of Reduction in Effective Capacities of High-Occupancy Vehicle Lanes Related to Traffic Behavior," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2065, pp. 47-53, 2008.
- [6] L. I. Zuyeva, "Equity issues in HOV-to-HOT conversion on I-85 North in Atlanta," Master of Science in the School of Civil and Environmental Engineering, City and Regional Planning, Georgia Institute of Technology, Atlanta, 2009.